DTU Library

Weighting and Aggregation in Life Cycle Assessment: Do Present Aggregated Single Scores Provide Correct Decision Support?

Kalbar, Pradip; Birkved, Morten; Nygaard, Simon Elsborg; Hauschild, Michael Zwicky

Published in:
Journal of Industrial Ecology

Link to article, DOI:
10.1111/jiec.12520

Publication date:
2017

Document Version
Peer reviewed version

Supplementary Information I

of

Weighting and Aggregation in LCA: Do Present Aggregated Single Scores Provide Correct Decision Support?

Pradip P. Kalbar[1*], Morten Birkved[1], Simon Elsborg Nygaard[2], and Michael Hauschild[1]

[1]Division for Quantitative Sustainability Assessment, DTU Management Engineering, Technical University of Denmark, Denmark.

[2]Department of Psychology and Behavioral Sciences, BSS, Aarhus University, Denmark

1. Methodology:

The PM-LCA model was applied to assess the outcome of the survey of the Danish residents (Kalbar et al. 2016). Using the survey, data related to housing, energy (heat and electricity), road transportation, air travel, food consumption, expenditures related to products and services, recycling habits and related sustainability behavior factors were collected. A total of 1281 respondents completed the questionnaire in its entirety. Out of this dataset only the first 1000 surveys were used for the present analysis, due to the software's limited capacity to handle larger datasets.

Yearly consumption patterns were estimated using the consumption-related data from the questionnaire. The consumption patterns were then assessed using the PM-LCA model. The reference house model was built in Gabi 6.0 (using the Ecoinvent 2.2 database), including production of all materials required for the construction of the reference house. Standard processes available in Gabi 6.0 (using the Ecoinvent 2.2 database) for heat, electricity, road transport (diesel and petrol cars with different Euro standards, public buses and trains) and air travel were used to quantify the impact potentials related to heat and electricity consumption as well as road transport and air travel. For estimation of the impact potentials related to food consumption, Simapro 8.0.4 (using the Ecoinvent 3.1 database) was used.

The ReCiPe 2008 impact assessment methodology Goedkoop et al. (2008) was used to estimate midpoint and endpoint impact potentials. The endpoints were further normalized using European normalization references Goedkoop et al. (2008). The normalized endpoint results were then weighted using different weighting schemes representing three cultural perspectives, *viz.*, hierarchical, individualistic and egalitarian Goedkoop et al. (2008). In addition, an equal weight scenario, as well as three extreme weighting schemes, were also used to quantify the weighting scheme’s impact on the single score. Table S1 summarizes the 7 weighting schemes applied in our comparison..

1.1 Dominance analysis using the Hasse Diagram Technique

The Hasse Diagram Technique (HDT) is a partial order ranking technique. Partial order techniques are non-compensatory approaches where no tradeoffs are allowed among the attributes and hence the MADM method exhibits no effect on the attribute values (Patil and Taillie 2004; R. Brüggemann, Schwaiger, and Negele 1995; Munda 2008). HDTs have been widely used in environmental decision making concerning the evaluation of water treatment technologies (Bick and Oron 2013), ecotoxicity tests (Brüggemann, Schwaiger, and Negele 1995), chemical substances (Brüggemann et al. 2006; Lerche et al. 2002), chemical ranking in LCA (Larsen et al. 2004), and water quality assessment (Voyslavov, Tsakovski, and Simeonov 2013).

In our study, the HDT was used to identify dominating respondents in the dataset. DART 2.05 (DART 2008) was used to obtain the preference level structure of the respondents. From the preference level structure, respondents dominating in all three endpoints could be easily identified and removed.

Table S1: Cultural perspective- and scenario-specific weighting schemes applied in the evaluation of single score aggregation methods

Perspective/ Scenario	Human health	Ecosystems	Resources	Total
Hierarchist	300	400	300	1000
Individualist	550	250	200	1000
Egalitarian	300	500	200	1000
Equal Weights	333.33	333.33	333.33	1000
Higher Weight to Human Health	800	100	100	1000
Higher Weight to Ecosystem	100	800	100	1000
Higher Weight to Resources	100	100	800	1000

2. Results:

Figure S1 shows the endpoint results of the LCA of the base dataset (n = 1000). As seen in Figure S1, these endpoints do not vary considerably. The endpoints for each of the respondents were used to generate ranks of the respondents using the Linear Weighted Sum (LWS) method of ReCiPe, (equation. 1 in main article). The same endpoint data were used to establish ranks using the TOPSIS method based on relative closeness (equation. 8 in main article). The best performing respondents (those having the best environmental profiles/lowest single score) were identified from this dataset using these two methods. The weighted normalized endpoint values of the best performing respondents are compared with PIS and shown in Figure S2. The PIS illustrated with red line triangles is the best possible environmental profile from the dataset; the objective of the methods used for obtaining single scores is to match the shape of the triangle formed by the PIS.

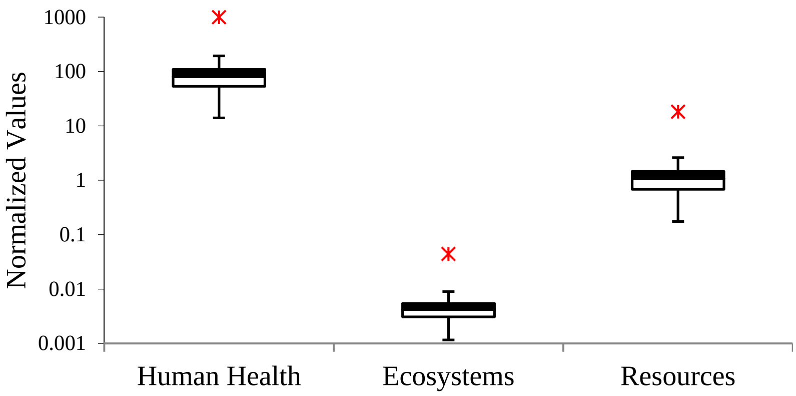


Figure S1: Variation of three endpoints (normalized values) obtained by assessment of the base datasets (n = 1000). The ends of the whisker are set at 1.5*IQR above the third quartile (Q3) and 1.5*IQR below the first quartile (Q1). The maximum values (outliers) are shown with red asterisk sign.

From the illustrations in Figure S2, it is clear that there is nearly complete agreement between TOPSIS and ReCiPe. The radar graphs show that the normalized endpoints plotted for the best performing respondents, as identified by the two methods, are identical (R678). However, it is also clear from Figure S2 that there is no change in rank results for various cultural perspectives (applying different weighting schemes) or when equal weights are used. To confirm this lack of rank sensitivity to weighting schemes, the Kendall's rank correlation coefficients (τ), between the ranks generated based on the single scores obtained using two methods for all of the perspectives considered, were estimated and are presented in Table S2. As seen in this table, the Kendall's Tau (τ) values for all sets of ranks are high (>0.9). In Figure 2 (main article), the first row of scatter plots shows the graphs of ranks generated by the two methods used for obtaining single scores. This confirms that ranks are not affected by the weighting procedure, regardless of methods used to obtain single scores (ReCiPe or TOPSIS). This is most likely due to the presence of dominating respondents in the dataset. To confirm this suspicion of the presence of dominating respondents, a dominance analysis using HDT was carried out. The results of HDT are provided in the Supplementary Information (SI) II. The dominance analysis revealed that there are dominating respondents present in the dataset, meaning that these respondents have high/low values for all three endpoints. Respondent R678 was identified as the best (environmentally) performing respondent by both of the methods used for obtaining single scores, regardless of cultural perspective. Using the results of dominance analysis (the level structure generated by HDT was used to reduce the dataset), 121 dominating respondents were removed from the dataset, and a new more homogenous (i.e., lacking the dominating respondents) dataset of aggregated respondents (n = 879) was created. Another reason for the dominance of R678 in all scenarios was the strong correlation of the three endpoints. As shown in Table S3, in the assessment of the base dataset, all three endpoints are strongly correlated ($\rho > 0.95$).

The boxplots in Figure S3 shows that the reduced dataset is now without outliers. The reduced dataset was further used to determine the ranks of respondents after using LWS and TOPSIS to obtain single scores. As in the initial analysis of the base dataset, the weighted normalized endpoint values of the best performing respondents were compared with the PIS and the results of this analysis is shown in Figure S4. As seen from the results in Figure S4, there is a disagreement between the two single score aggregation methods (LWS and TOPSIS) in the hierarchical and individualist perspectives as well as equal weights scenarios. This shows that after removing the dominating respondents, the method used to obtain single scores does now affect the identification of the best performing respondent. However, the results also show that the weighting scheme has no effect on the ranks generated by the two methods, as the best performing respondents remain the same in each of the scenarios except for the egalitarian scenario. The low influence of the weighting scheme on the reduced dataset is also evident from the strong rank correlations (see Table S4), highlighting the fact that cultural perspectives still have very little effect on the ranks.

The reason for the lack of influence of the weighting scheme on the ranks is once again attributed to the strong correlation among the endpoint dataset (see Table S3). The strong endpoint correlation observed in the dataset (ρ = 0.95) affects the results of the ranking. The results reveal that when the endpoints are strongly correlated, the weighting of the individual endpoints in relation to their aggregation into a single score is low. This lack of influence of the weighting not only highlights the need for assessing the weighted endpoint aggregation, but also the need to assess the actual ranking methods.

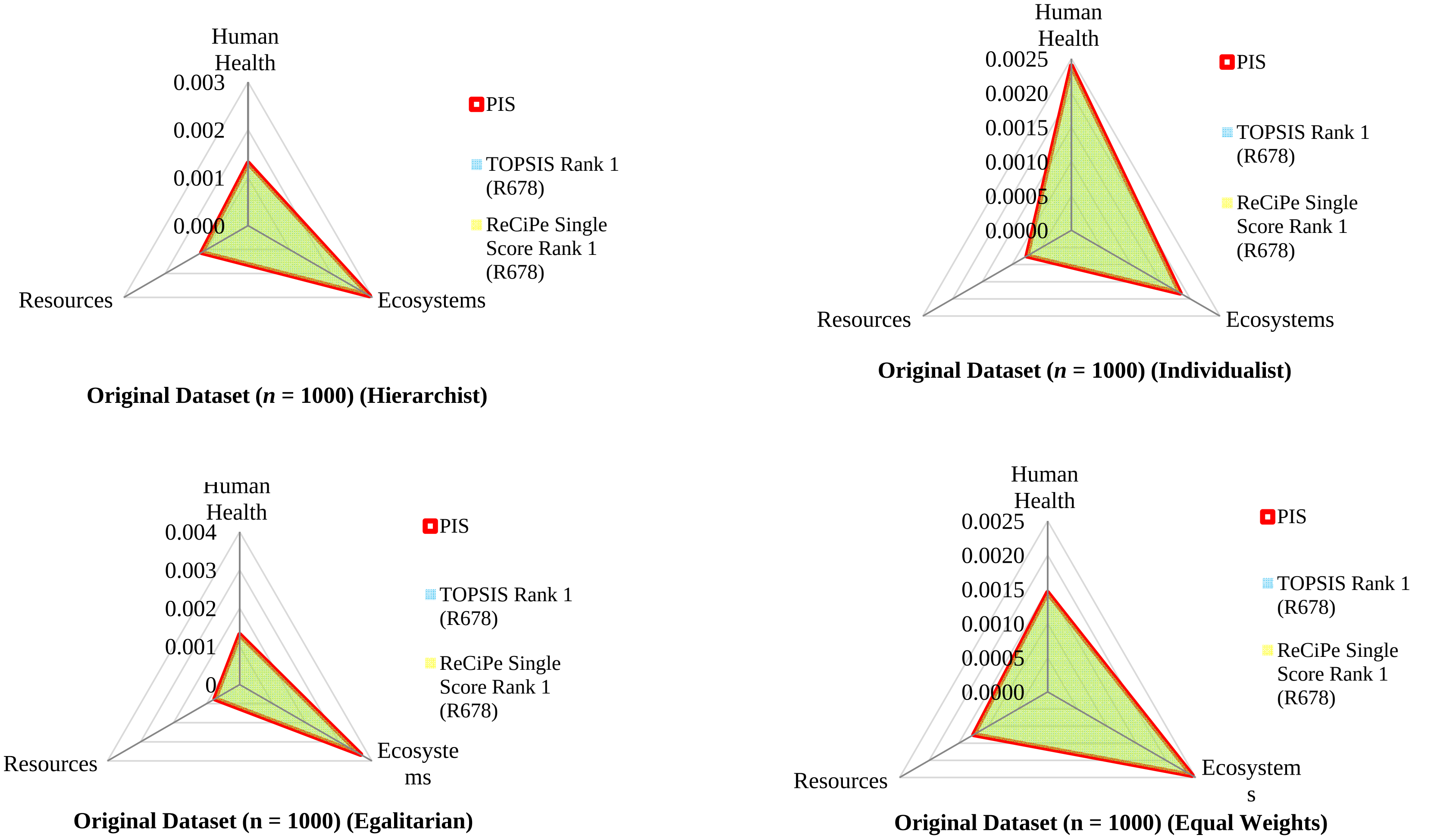


Figure S2: The radar plot shows the weighted normalized values of endpoints for best performing respondents (out of the base dataset) for various cultural perspectives and the equal weight scenario

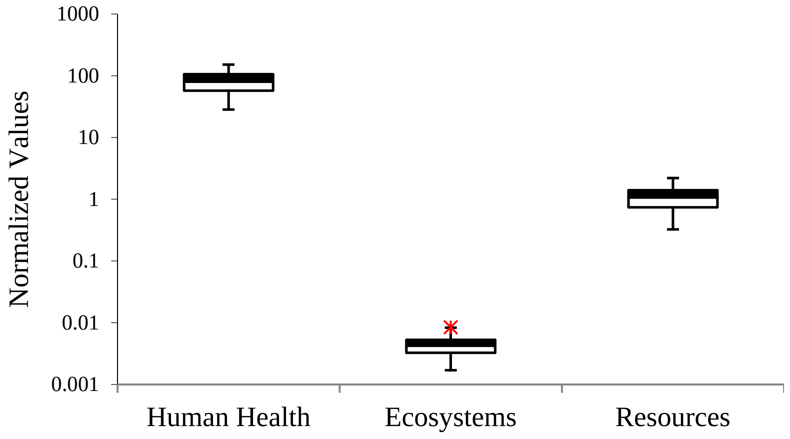


Figure S3: Variation of three end points (normalized values) in reduced data sets after dominance analysis (n = 879)

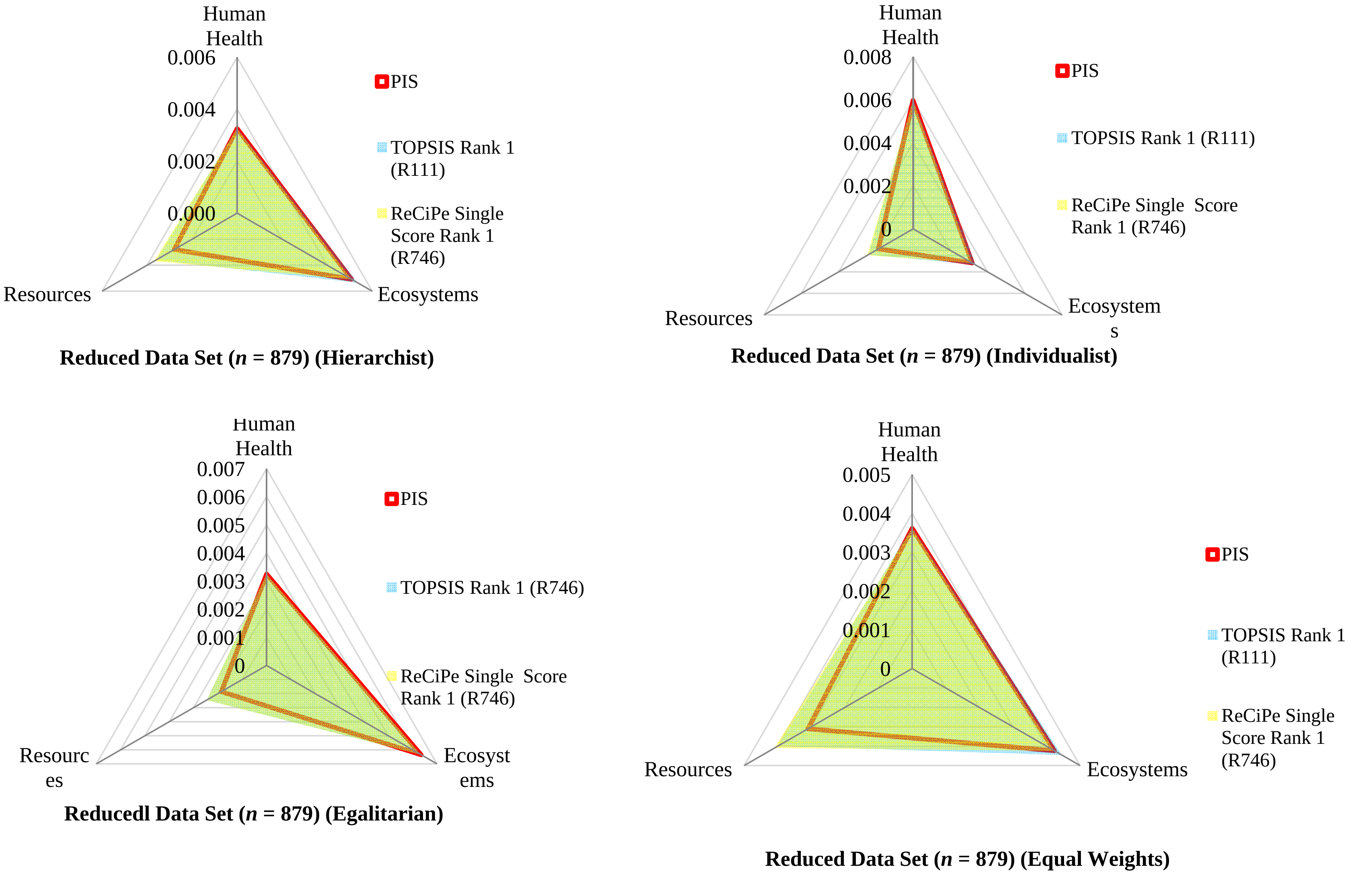


Figure S4: The radar plot shows the weighted normalized values of endpoints for best performing respondents of the reduced data set for various cultural perspectives and equal weight scenario

Table S2: Results of Kendall's rank correlation coefficient (τ) between the ranks generated by the two methods for various perspectives (Base dataset)

	Hierarchist - TOPSIS Ranks	Individualist - TOPSIS Ranks	Egalitarian - TOPSIS Ranks	Equal Weights - TOPSIS Ranks	Hierarchist - ReCiPe Single Score Ranks	Individualist - ReCiPe Single Score Ranks	Egalitarian - ReCiPe Single Score Ranks	Equal Weights - ReCiPe Single Score Ranks
Hierarchist - TOPSIS Ranks	**1.00**	**0.94**	**0.96**	**0.98**	**0.92**	**0.92**	**0.92**	**0.92**
Individualist - TOPSIS Ranks	**0.94**	**1.00**	**0.92**	**0.96**	**0.98**	**0.98**	**0.98**	**0.98**
Egalitarian - TOPSIS Ranks	**0.96**	**0.92**	**1.00**	**0.94**	0.89	0.89	0.89	0.89
Equal Weights - TOPSIS Ranks	**0.98**	**0.96**	**0.94**	**1.00**	**0.94**	**0.94**	**0.94**	**0.94**
Hierarchist - ReCiPe Single Score Ranks	**0.92**	**0.98**	0.89	**0.94**	**1.00**	**1.00**	**1.00**	**1.00**
Individualist - ReCiPe Single Score Ranks	**0.92**	**0.98**	0.89	**0.94**	**1.00**	**1.00**	**1.00**	**1.00**
Egalitarian - ReCiPe Single Score Ranks	**0.92**	**0.98**	0.89	**0.94**	**1.00**	**1.00**	**1.00**	**1.00**
Equal Weights - ReCiPe Single Score Ranks	**0.92**	**0.98**	0.89	**0.94**	**1.00**	**1.00**	**1.00**	**1.00**

Table S3: Results of the Sperman's rank correlation coefficient for three datasets used for analysis (grey shaded correlation coefficient values are not signficant at 0.05 level). The endpoints are derived using hirachical perspective.

Original Data Set (*n* = 1000)				**Reduced Data Set after Dominance Analysis (*n* = 879)**				**Random Data Set (*n* = 879)**			
	Human Health	Ecosystems	Resources		Human Health	Ecosystems	Resources		Human Health	Ecosystems	Resources
Human Health	1.00	0.96	0.98	Human Health	1.00	0.94	0.97	Human Health	1.00	-0.02	-0.02
Ecosystems	--	1.00	0.94	Ecosystems	--	1.00	0.91	Ecosystems	--	1.00	0.00
Resources	--	--	1.00	Resources	--	--	1.00	Resources	--	--	1.00

Table S4: Results of Kendall's rank collereation coefficienct (τ) between the ranks generated by the two methods for various perspectives (reduced Data Set)

	Hierarchist - TOPSIS Ranks	Individualist - TOPSIS Ranks	Egalitarian - TOPSIS Ranks	Equal Weights - TOPSIS Ranks	Hierarchist - ReCiPe Single Score Ranks	Individualist - ReCiPe Single Score Ranks	Egalitarian - ReCiPe Single Score Ranks	Equal Weights - ReCiPe Single Score Ranks
Hierarchist - TOPSIS Ranks	**1.00**	**0.93**	**0.95**	**0.97**	**0.90**	**0.90**	**0.90**	**0.90**
Individualist - TOPSIS Ranks	**0.93**	**1.00**	**0.90**	**0.95**	**0.96**	**0.96**	**0.96**	**0.96**
Egalitarian - TOPSIS Ranks	**0.95**	**0.90**	**1.00**	**0.92**	**0.86**	**0.86**	**0.86**	**0.86**
Equal Weights - TOPSIS Ranks	**0.97**	**0.95**	**0.92**	**1.00**	**0.92**	**0.92**	**0.92**	**0.92**
Hierarchist - ReCiPe Single Score Ranks	**0.90**	**0.96**	**0.86**	**0.92**	**1.00**	**1.00**	**1.00**	**1.00**
Individualist - ReCiPe Single Score Ranks	**0.90**	**0.96**	**0.86**	**0.92**	**1.00**	**1.00**	**1.00**	**1.00**
Egalitarian - ReCiPe Single Score Ranks	**0.90**	**0.96**	**0.86**	**0.92**	**1.00**	**1.00**	**1.00**	**1.00**
Equal Weights - ReCiPe Single Score Ranks	**0.90**	**0.96**	**0.86**	**0.92**	**1.00**	**1.00**	**1.00**	**1.00**

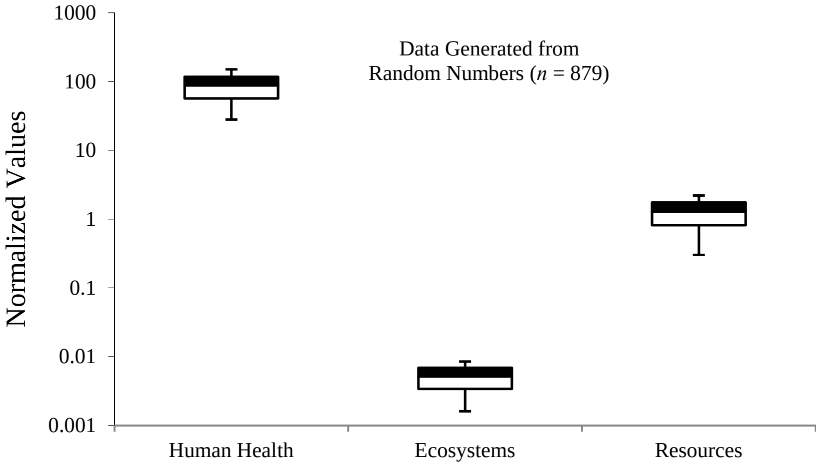
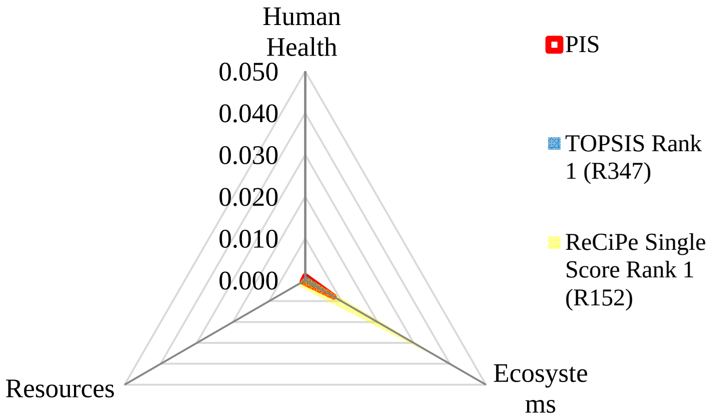
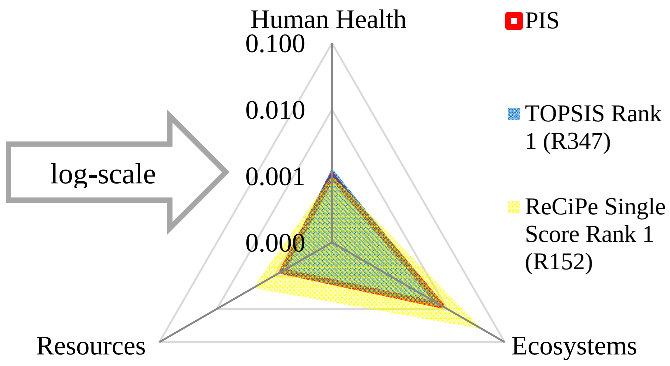


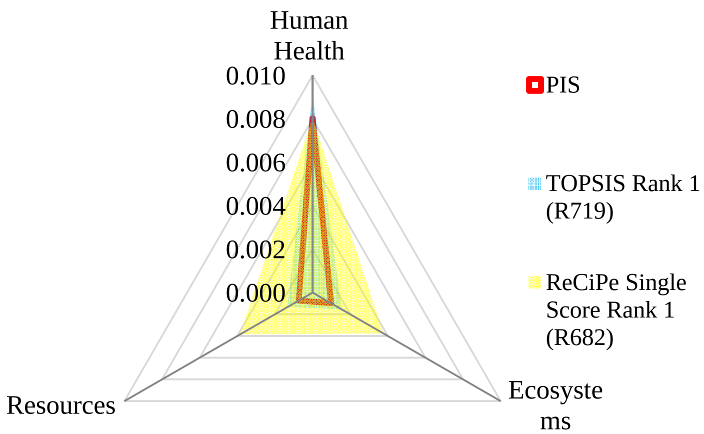
Figure S5: Variation of three end points (normalized values) in randomly generated data sets (*n* = 879)



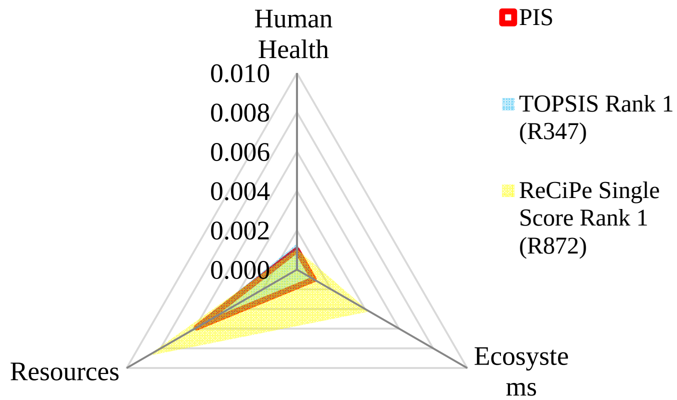
Random Data Set (*n* = 879) Very High Weight to Ecosystem



Random Data Set (*n* = 879) Very High Weight to Ecosystem



Random Data Set (*n* = 879) Very High Weight to Human Health



Random Data Set (*n* = 879) Very High Weight to Resources

Figure S6: The radar plot shows the weighted normalized values of endpoints for best performing respondents of the random data set for various cultural weighting schemes

Table S5: Results of Kendall's rank collereation coefficienct (τ) between the ranks generated by the two methods for various extreme weights scenarios (Random Data Set)

	High weight to Ecosystem - TOPSIS Ranks	High Weight to Human Health- TOPSIS Ranks	High Weight to Resources - TOPSIS Ranks	High weight to Ecosystem - ReCiPe Single Score Ranks	High Weight to Human Health - ReCiPe Single Score Ranks	High Weight to Resources - ReCiPe Single Score Ranks
High weight to Ecosystem - TOPSIS Ranks	1.00	0.04	0.05	0.01	0.01	0.01
High Weight to Human Health- TOPSIS Ranks	0.04	1.00	0.04	0.97	0.96	0.94
High Weight to Resources - TOPSIS Ranks	0.05	0.04	1.00	0.02	0.01	0.09
High weight to Ecosystem - ReCiPe Single Score Ranks	0.01	0.97	0.02	1.00	0.99	0.93
High Weight to Human Health - ReCiPe Single Score Ranks	0.01	0.96	0.01	0.99	1.00	0.92
High Weight to Resources - ReCiPe Single Score Ranks	0.01	0.94	0.09	0.93	0.92	1.00

References:

Bick, Amos, and Gideon Oron. 2013. "Boron Removal from Seawater Reverse Osmosis Permeate: A Hasse Diagram Analysis of Current Technologies." *Desalination* 310. Elsevier B.V.: 34–38. doi:10.1016/j.desal.2012.10.002.

Brüggemann, R., J. Schwaiger, and R.D. Negele. 1995. "Applying Hasse Diagram Technique for the Evaluation of Toxicological Fish Tests." *Chemosphere* 30 (9): 1767–80. doi:10.1016/0045-6535(95)00061-C.

Brüggemann, Rainer, Lars Carlsen, Dorte B. Lerche, and Peter B. Sorensen. 2006. "A Comparison of Partial Order Technique with Three Methods of Multi-Criteria Analysis for Ranking of Chemical Substance." *Partial Order in Environmental Sciences and Chemistry*, 237–56. doi:10.1007/3-540-33970-1_10.

DART. 2008. "DART 2.05 (Decision Analysis by Ranking Techniques)." https://eurl-ecvam.jrc.ec.europa.eu/laboratories-research/predictive_toxicology/qsar_tools/DART.

Goedkoop, Mark, Reinout Heijungs, Mark a. J. Huijbregts, An De Schryver, Jaap Struijs, and Rosalie Van Zelm. 2008. "ReCiPe 2008, A Life Cycle Impact Assessment Method Which Comprises Harminised Category Indicators at the Midpoint and the Endpoint Level, First Edition, Report I: Characterisation." *Available from Internet: Http://www. Lcia-Recipe. Net*, 126. http://www.lcia-recipe.net/.

Kalbar, Pradip P., Morten Birkved, Simon Kabins, and Simon E Nygaard. 2016. "Personal-Metabolism (PM) Coupled with Life Cycle Assessment (LCA) Model: Danish Case Study - Accepted In Press." *Environment International*. doi:10.1016/j.envint.2016.02.032.

Larsen, Henrik Fred, Morten Birkved, Michael Hauschild, David W. Pennington, and Jeroen B. Guinée. 2004. "Evaluation of Selection Methods for Toxicological Impacts in LCA Recommendations for OMNIITOX." *The International Journal of Life Cycle Assessment* 9 (5): 307–19. doi:10.1007/BF02979420.

Lerche, D., E. Van de Plassche, a. Schwegler, and F. Balk. 2002. "Selecting Chemical Substances for the UN-ECE POP Protocol." *Chemosphere* 47 (6): 617–30. doi:10.1016/S0045-6535(02)00028-0.

Munda, Giuseppe. 2008. *Social Evaluation for a Sustainable Economy*. Springer– Verlag, Berlin.

Patil, G. P., and C. Taillie. 2004. "Multiple Indicators, Partially Ordered Sets, and Linear Extensions: Multi-Criterion Ranking and Prioritization." *Environmental and Ecological Statistics* 11 (Mcmc): 199–228. doi:10.1023/B:EEST.0000027209.93218.d9.

Voyslavov, Tsvetomil, Stefan Tsakovski, and Vasil Simeonov. 2013. "Hasse Diagram Technique as a Tool for Water Quality Assessment." *Analytica Chimica Acta* 770. Elsevier B.V.: 29–35. doi:10.1016/j.aca.2013.01.063.

Supplementary Information II

of

Weighting and Aggregation in LCA: Does Present Aggregated Single Scores Provide Correct Decision Support?

Pradip P. Kalbar[1*], Morten Birkved[1], Simon Elsborg Nygaard[2], and Michael Hauschild[1]

[1]Division for Quantitative Sustainability Assessment, DTU Management Engineering, Technical University of Denmark, Denmark.

[2]Department of Psychology and Behavioural Sciences, BSS, Aarhus University, Denmark

Partial Ranking (Hasse Diagram Technique) Report Using Dart 2.05

No. of criteria 3
No. of objects 1000
No. of levels (NL) 165
No. of elements in the largest level (NEL) 15
Comparability (V(N)) 465990
Contradictions (U(N)) 33510
No. of equivalence classes (Z) 1000
No. of equivalence classes with more than one obj (NECA) 0
No. of maximals (NMax) 1
Maximal elements: Obj.730
No. of minimals (NMin) 1
Minimal elements: Obj.678
No. of isolated (NIso) 0

Level structure:
Level 165 (1 elements): Obj.730
Level 164 (1 elements): Obj.729
Level 163 (2 elements): Obj.600; Obj.700
Level 162 (1 elements): Obj.57
Level 161 (2 elements): Obj.249; Obj.834
Level 160 (3 elements): Obj.237; Obj.898; Obj.919
Level 159 (3 elements): Obj.521; Obj.756; Obj.774
Level 158 (3 elements): Obj.550; Obj.575; Obj.933
Level 157 (2 elements): Obj.325; Obj.423
Level 156 (2 elements): Obj.478; Obj.503

Level 155 (2 elements): Obj.602; Obj.949
Level 154 (4 elements): Obj.62; Obj.240; Obj.960; Obj.964
Level 153 (6 elements): Obj.167; Obj.171; Obj.205; Obj.398; Obj.646; Obj.1000
Level 152 (1 elements): Obj.944
Level 151 (1 elements): Obj.280
Level 150 (4 elements): Obj.139; Obj.163; Obj.393; Obj.640
Level 149 (5 elements): Obj.135; Obj.278; Obj.359; Obj.567; Obj.997
Level 148 (4 elements): Obj.30; Obj.588; Obj.930; Obj.996
Level 147 (5 elements): Obj.273; Obj.334; Obj.431; Obj.563; Obj.982
Level 146 (4 elements): Obj.480; Obj.796; Obj.961; Obj.991
Level 145 (3 elements): Obj.332; Obj.863; Obj.897
Level 144 (6 elements): Obj.145; Obj.276; Obj.426; Obj.537; Obj.896; Obj.918
Level 143 (6 elements): Obj.303; Obj.327; Obj.705; Obj.807; Obj.943; Obj.979
Level 142 (5 elements): Obj.150; Obj.285; Obj.648; Obj.797; Obj.998
Level 141 (8 elements): Obj.322; Obj.429; Obj.512; Obj.598; Obj.613; Obj.745; Obj.970; Obj.973
Level 140 (8 elements): Obj.289; Obj.387; Obj.473; Obj.514; Obj.519; Obj.591; Obj.636; Obj.971
Level 139 (5 elements): Obj.351; Obj.732; Obj.783; Obj.940; Obj.975
Level 138 (11 elements): Obj.80; Obj.136; Obj.161; Obj.168; Obj.169; Obj.189; Obj.281; Obj.432; Obj.505; Obj.552; Obj.764
Level 137 (7 elements): Obj.34; Obj.151; Obj.190; Obj.293; Obj.499; Obj.679; Obj.934
Level 136 (7 elements): Obj.314; Obj.472; Obj.553; Obj.589; Obj.618; Obj.734; Obj.967
Level 135 (8 elements): Obj.252; Obj.479; Obj.544; Obj.622; Obj.744; Obj.917; Obj.946; Obj.990
Level 134 (6 elements): Obj.173; Obj.396; Obj.421; Obj.612; Obj.851; Obj.958
Level 133 (7 elements): Obj.286; Obj.313; Obj.427; Obj.513; Obj.562; Obj.728; Obj.817
Level 132 (7 elements): Obj.152; Obj.390; Obj.621; Obj.623; Obj.688; Obj.769; Obj.900
Level 131 (6 elements): Obj.180; Obj.231; Obj.418; Obj.490; Obj.815; Obj.981
Level 130 (9 elements): Obj.147; Obj.157; Obj.170; Obj.203; Obj.287; Obj.328; Obj.339; Obj.546; Obj.989
Level 129 (8 elements): Obj.206; Obj.282; Obj.498; Obj.517; Obj.607; Obj.666; Obj.727; Obj.966
Level 128 (8 elements): Obj.49; Obj.197; Obj.264; Obj.302; Obj.395; Obj.441; Obj.573; Obj.925
Level 127 (7 elements): Obj.40; Obj.232; Obj.336; Obj.433; Obj.565; Obj.707; Obj.920
Level 126 (7 elements): Obj.142; Obj.200; Obj.290; Obj.333; Obj.531; Obj.872; Obj.932
Level 125 (4 elements): Obj.175; Obj.438; Obj.529; Obj.910
Level 124 (9 elements): Obj.191; Obj.308; Obj.319; Obj.324; Obj.361; Obj.410; Obj.492; Obj.701; Obj.704
Level 123 (6 elements): Obj.279; Obj.557; Obj.585; Obj.593; Obj.703; Obj.935
Level 122 (6 elements): Obj.176; Obj.186; Obj.227; Obj.402; Obj.495; Obj.597
Level 121 (9 elements): Obj.105; Obj.196; Obj.213; Obj.261; Obj.500; Obj.614; Obj.650; Obj.800; Obj.924
Level 120 (7 elements): Obj.153; Obj.160; Obj.508; Obj.559; Obj.755; Obj.868; Obj.988
Level 119 (7 elements): Obj.294; Obj.355; Obj.425; Obj.436; Obj.470; Obj.539; Obj.541
Level 118 (9 elements): Obj.16; Obj.43; Obj.207; Obj.373; Obj.556; Obj.641; Obj.931; Obj.936; Obj.976
Level 117 (9 elements): Obj.210; Obj.241; Obj.416; Obj.604; Obj.642; Obj.788; Obj.809; Obj.942; Obj.956
Level 116 (5 elements): Obj.202; Obj.251; Obj.338; Obj.348; Obj.555
Level 115 (7 elements): Obj.162; Obj.178; Obj.215; Obj.430; Obj.518; Obj.572; Obj.880
Level 114 (12 elements): Obj.7; Obj.229; Obj.309; Obj.392; Obj.469; Obj.489; Obj.583; Obj.719; Obj.739; Obj.794; Obj.812; Obj.987
Level 113 (8 elements): Obj.349; Obj.352; Obj.401; Obj.515; Obj.893; Obj.923; Obj.945; Obj.999
Level 112 (9 elements): Obj.216; Obj.258; Obj.284; Obj.357; Obj.566; Obj.798; Obj.875; Obj.885; Obj.922

Level 111 (11 elements): Obj.181; Obj.201; Obj.246; Obj.250; Obj.548; Obj.560; Obj.627; Obj.632; Obj.693; Obj.808; Obj.969

Level 110 (9 elements): Obj.164; Obj.172; Obj.228; Obj.360; Obj.811; Obj.879; Obj.887; Obj.941; Obj.948

Level 109 (10 elements): Obj.248; Obj.275; Obj.299; Obj.331; Obj.376; Obj.551; Obj.586; Obj.742; Obj.775; Obj.947

Level 108 (8 elements): Obj.154; Obj.388; Obj.501; Obj.601; Obj.645; Obj.790; Obj.801; Obj.950

Level 107 (11 elements): Obj.140; Obj.320; Obj.374; Obj.420; Obj.434; Obj.542; Obj.568; Obj.576; Obj.584; Obj.686; Obj.717

Level 106 (11 elements): Obj.127; Obj.138; Obj.174; Obj.230; Obj.463; Obj.484; Obj.486; Obj.757; Obj.822; Obj.837; Obj.951

Level 105 (15 elements): Obj.235; Obj.238; Obj.262; Obj.311; Obj.350; Obj.379; Obj.391; Obj.457; Obj.538; Obj.652; Obj.806; Obj.840; Obj.870; Obj.928; Obj.984

Level 104 (11 elements): Obj.110; Obj.133; Obj.209; Obj.291; Obj.404; Obj.483; Obj.549; Obj.638; Obj.709; Obj.748; Obj.876

Level 103 (9 elements): Obj.155; Obj.179; Obj.292; Obj.335; Obj.337; Obj.409; Obj.511; Obj.741; Obj.980

Level 102 (7 elements): Obj.12; Obj.267; Obj.422; Obj.475; Obj.525; Obj.763; Obj.939

Level 101 (12 elements): Obj.259; Obj.307; Obj.406; Obj.467; Obj.497; Obj.504; Obj.516; Obj.653; Obj.750; Obj.765; Obj.773; Obj.818

Level 100 (10 elements): Obj.208; Obj.263; Obj.323; Obj.340; Obj.344; Obj.502; Obj.628; Obj.706; Obj.708; Obj.978

Level 99 (7 elements): Obj.8; Obj.82; Obj.399; Obj.534; Obj.736; Obj.889; Obj.921

Level 98 (10 elements): Obj.113; Obj.192; Obj.204; Obj.242; Obj.596; Obj.629; Obj.651; Obj.836; Obj.915; Obj.972

Level 97 (12 elements): Obj.11; Obj.146; Obj.183; Obj.188; Obj.234; Obj.298; Obj.346; Obj.415; Obj.428; Obj.466; Obj.810; Obj.866

Level 96 (12 elements): Obj.124; Obj.149; Obj.187; Obj.268; Obj.288; Obj.371; Obj.381; Obj.570; Obj.590; Obj.671; Obj.766; Obj.954

Level 95 (10 elements): Obj.9; Obj.66; Obj.134; Obj.185; Obj.220; Obj.510; Obj.649; Obj.953; Obj.962; Obj.965

Level 94 (6 elements): Obj.29; Obj.247; Obj.260; Obj.564; Obj.792; Obj.895

Level 93 (7 elements): Obj.37; Obj.305; Obj.342; Obj.403; Obj.408; Obj.733; Obj.882

Level 92 (7 elements): Obj.60; Obj.219; Obj.523; Obj.536; Obj.619; Obj.720; Obj.841

Level 91 (6 elements): Obj.96; Obj.265; Obj.582; Obj.647; Obj.778; Obj.805

Level 90 (7 elements): Obj.55; Obj.414; Obj.496; Obj.610; Obj.858; Obj.909; Obj.952

Level 89 (6 elements): Obj.417; Obj.569; Obj.630; Obj.838; Obj.842; Obj.908

Level 88 (6 elements): Obj.148; Obj.326; Obj.471; Obj.749; Obj.890; Obj.974

Level 87 (5 elements): Obj.405; Obj.464; Obj.574; Obj.692; Obj.820

Level 86 (9 elements): Obj.14; Obj.158; Obj.184; Obj.269; Obj.458; Obj.547; Obj.881; Obj.937; Obj.977

Level 85 (8 elements): Obj.36; Obj.141; Obj.343; Obj.358; Obj.375; Obj.437; Obj.558; Obj.587

Level 84 (6 elements): Obj.42; Obj.243; Obj.321; Obj.634; Obj.955; Obj.985

Level 83 (6 elements): Obj.540; Obj.672; Obj.673; Obj.699; Obj.867; Obj.914

Level 82 (8 elements): Obj.116; Obj.347; Obj.382; Obj.561; Obj.594; Obj.615; Obj.680; Obj.846

Level 81 (7 elements): Obj.69; Obj.118; Obj.177; Obj.253; Obj.477; Obj.869; Obj.957

Level 80 (10 elements): Obj.25; Obj.47; Obj.120; Obj.143; Obj.520; Obj.543; Obj.616; Obj.639; Obj.690; Obj.983

Level 79 (10 elements): Obj.144; Obj.255; Obj.317; Obj.440; Obj.687; Obj.743; Obj.762; Obj.901; Obj.903; Obj.995

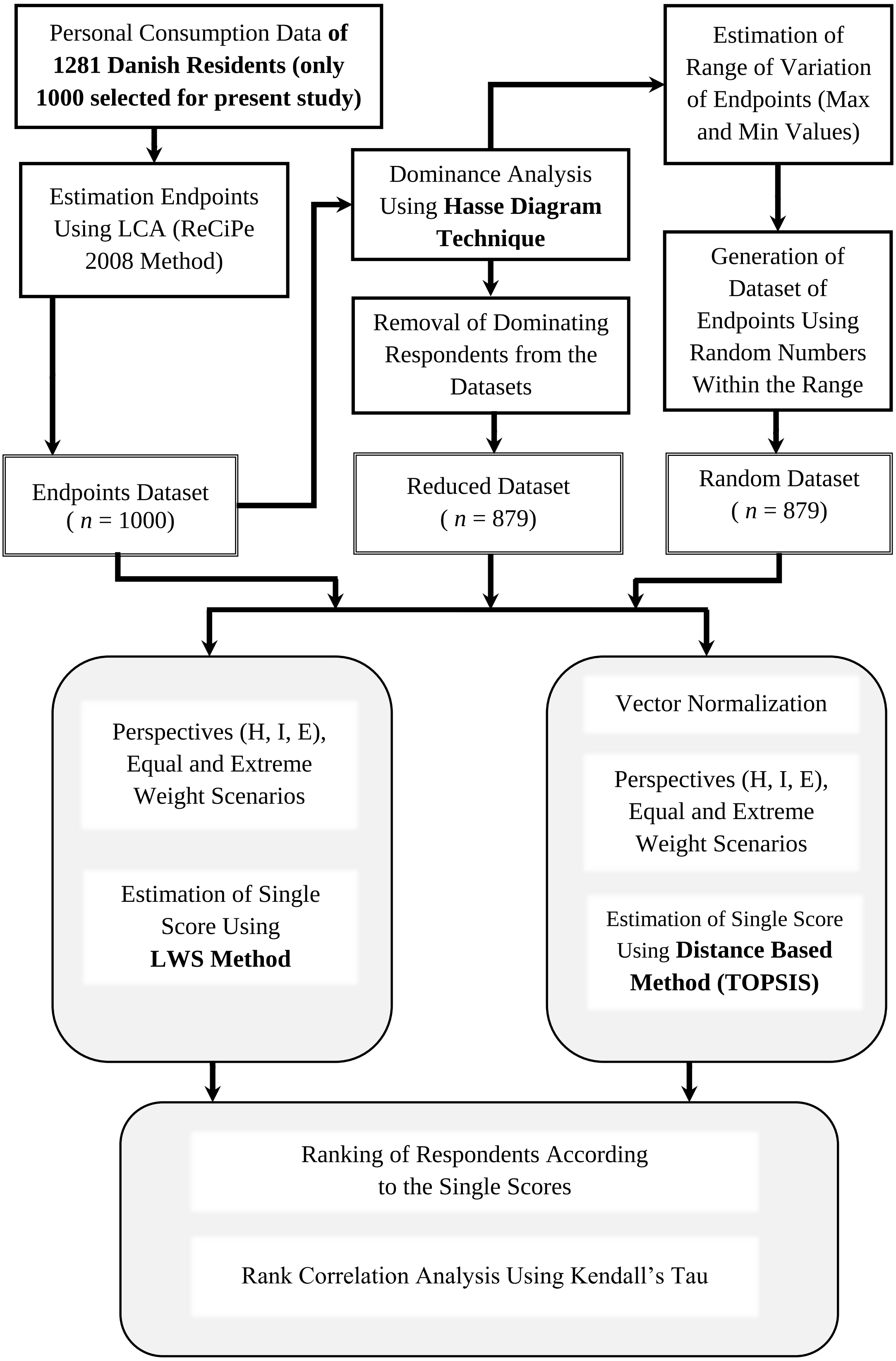
Level 78 (6 elements): Obj.329; Obj.341; Obj.354; Obj.439; Obj.599; Obj.710

Level 77 (4 elements): Obj.254; Obj.318; Obj.474; Obj.507

Level 76 (10 elements): Obj.18; Obj.28; Obj.223; Obj.372; Obj.380; Obj.571; Obj.592; Obj.691; Obj.724; Obj.738
Level 75 (9 elements): Obj.156; Obj.386; Obj.468; Obj.485; Obj.670; Obj.698; Obj.802; Obj.902; Obj.963
Level 74 (6 elements): Obj.75; Obj.444; Obj.581; Obj.871; Obj.904; Obj.938
Level 73 (9 elements): Obj.61; Obj.226; Obj.435; Obj.522; Obj.625; Obj.662; Obj.776; Obj.781; Obj.856
Level 72 (8 elements): Obj.193; Obj.239; Obj.368; Obj.491; Obj.524; Obj.526; Obj.644; Obj.723
Level 71 (6 elements): Obj.378; Obj.412; Obj.413; Obj.726; Obj.862; Obj.986
Level 70 (9 elements): Obj.274; Obj.277; Obj.528; Obj.579; Obj.595; Obj.722; Obj.740; Obj.833; Obj.905
Level 69 (8 elements): Obj.35; Obj.270; Obj.383; Obj.384; Obj.419; Obj.580; Obj.633; Obj.635
Level 68 (7 elements): Obj.56; Obj.68; Obj.609; Obj.611; Obj.674; Obj.761; Obj.804
Level 67 (5 elements): Obj.385; Obj.482; Obj.603; Obj.894; Obj.959
Level 66 (6 elements): Obj.244; Obj.345; Obj.493; Obj.620; Obj.793; Obj.799
Level 65 (6 elements): Obj.13; Obj.533; Obj.716; Obj.735; Obj.848; Obj.873
Level 64 (7 elements): Obj.48; Obj.53; Obj.236; Obj.297; Obj.506; Obj.545; Obj.656
Level 63 (8 elements): Obj.17; Obj.71; Obj.100; Obj.103; Obj.126; Obj.182; Obj.411; Obj.737
Level 62 (9 elements): Obj.21; Obj.84; Obj.117; Obj.212; Obj.218; Obj.365; Obj.657; Obj.874; Obj.911
Level 61 (5 elements): Obj.316; Obj.476; Obj.758; Obj.823; Obj.926
Level 60 (4 elements): Obj.300; Obj.400; Obj.554; Obj.913
Level 59 (7 elements): Obj.211; Obj.245; Obj.445; Obj.451; Obj.665; Obj.731; Obj.993
Level 58 (6 elements): Obj.194; Obj.222; Obj.682; Obj.683; Obj.751; Obj.855
Level 57 (5 elements): Obj.38; Obj.72; Obj.165; Obj.676; Obj.725
Level 56 (4 elements): Obj.88; Obj.98; Obj.481; Obj.527
Level 55 (6 elements): Obj.51; Obj.447; Obj.454; Obj.535; Obj.660; Obj.789
Level 54 (5 elements): Obj.92; Obj.214; Obj.225; Obj.455; Obj.605
Level 53 (7 elements): Obj.131; Obj.166; Obj.271; Obj.363; Obj.394; Obj.532; Obj.877
Level 52 (6 elements): Obj.109; Obj.330; Obj.465; Obj.509; Obj.617; Obj.695
Level 51 (7 elements): Obj.3; Obj.77; Obj.83; Obj.114; Obj.577; Obj.631; Obj.849
Level 50 (5 elements): Obj.5; Obj.15; Obj.52; Obj.65; Obj.843
Level 49 (5 elements): Obj.39; Obj.369; Obj.664; Obj.715; Obj.865
Level 48 (8 elements): Obj.159; Obj.667; Obj.668; Obj.685; Obj.702; Obj.825; Obj.892; Obj.994
Level 47 (8 elements): Obj.2; Obj.26; Obj.78; Obj.304; Obj.362; Obj.446; Obj.487; Obj.779
Level 46 (4 elements): Obj.104; Obj.782; Obj.814; Obj.883
Level 45 (6 elements): Obj.32; Obj.54; Obj.97; Obj.217; Obj.626; Obj.714
Level 44 (6 elements): Obj.20; Obj.108; Obj.366; Obj.624; Obj.819; Obj.821
Level 43 (4 elements): Obj.130; Obj.283; Obj.456; Obj.711
Level 42 (6 elements): Obj.58; Obj.102; Obj.233; Obj.257; Obj.377; Obj.795
Level 41 (5 elements): Obj.364; Obj.816; Obj.826; Obj.831; Obj.886
Level 40 (5 elements): Obj.353; Obj.768; Obj.828; Obj.845; Obj.854
Level 39 (4 elements): Obj.137; Obj.312; Obj.718; Obj.968
Level 38 (5 elements): Obj.272; Obj.306; Obj.310; Obj.494; Obj.813
Level 37 (10 elements): Obj.63; Obj.70; Obj.86; Obj.442; Obj.450; Obj.453; Obj.578; Obj.713; Obj.747; Obj.803
Level 36 (4 elements): Obj.266; Obj.661; Obj.770; Obj.878
Level 35 (6 elements): Obj.44; Obj.119; Obj.199; Obj.370; Obj.424; Obj.827
Level 34 (6 elements): Obj.19; Obj.448; Obj.452; Obj.608; Obj.857; Obj.899
Level 33 (5 elements): Obj.31; Obj.59; Obj.73; Obj.530; Obj.643

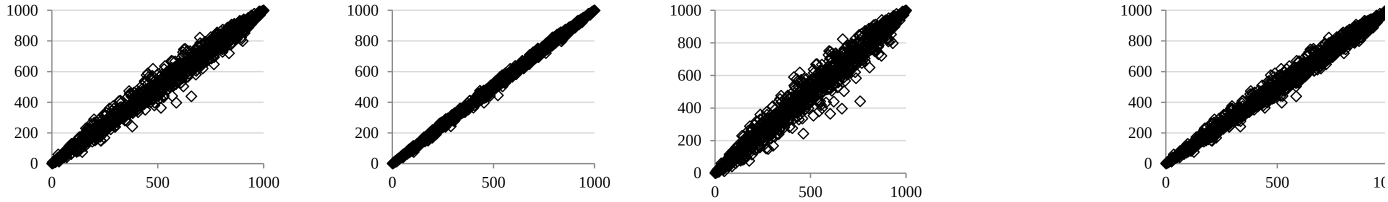
Level 32 (4 elements):	Obj.129; Obj.315; Obj.677; Obj.861
Level 31 (5 elements):	Obj.85; Obj.256; Obj.663; Obj.754; Obj.891
Level 30 (4 elements):	Obj.367; Obj.681; Obj.785; Obj.844
Level 29 (5 elements):	Obj.389; Obj.449; Obj.684; Obj.712; Obj.853
Level 28 (5 elements):	Obj.4; Obj.41; Obj.637; Obj.777; Obj.992
Level 27 (5 elements):	Obj.296; Obj.356; Obj.407; Obj.659; Obj.772
Level 26 (4 elements):	Obj.64; Obj.112; Obj.780; Obj.916
Level 25 (4 elements):	Obj.101; Obj.198; Obj.295; Obj.824
Level 24 (9 elements):	Obj.10; Obj.27; Obj.46; Obj.115; Obj.121; Obj.655; Obj.696; Obj.753; Obj.760
Level 23 (5 elements):	Obj.45; Obj.91; Obj.488; Obj.829; Obj.864
Level 22 (5 elements):	Obj.22; Obj.221; Obj.689; Obj.746; Obj.906
Level 21 (5 elements):	Obj.93; Obj.111; Obj.195; Obj.301; Obj.771
Level 20 (5 elements):	Obj.87; Obj.224; Obj.791; Obj.888; Obj.912
Level 19 (3 elements):	Obj.50; Obj.675; Obj.929
Level 18 (5 elements):	Obj.89; Obj.90; Obj.658; Obj.835; Obj.847
Level 17 (5 elements):	Obj.694; Obj.759; Obj.767; Obj.850; Obj.852
Level 16 (4 elements):	Obj.122; Obj.132; Obj.461; Obj.830
Level 15 (4 elements):	Obj.95; Obj.128; Obj.832; Obj.839
Level 14 (5 elements):	Obj.24; Obj.33; Obj.459; Obj.460; Obj.669
Level 13 (4 elements):	Obj.74; Obj.654; Obj.784; Obj.927
Level 12 (4 elements):	Obj.67; Obj.106; Obj.123; Obj.697
Level 11 (2 elements):	Obj.752; Obj.787
Level 10 (3 elements):	Obj.1; Obj.6; Obj.81
Level 9 (3 elements):	Obj.23; Obj.99; Obj.443
Level 8 (1 elements):	Obj.907
Level 7 (3 elements):	Obj.786; Obj.859; Obj.860
Level 6 (4 elements):	Obj.94; Obj.397; Obj.462; Obj.884
Level 5 (2 elements):	Obj.125; Obj.606
Level 4 (1 elements):	Obj.107
Level 3 (1 elements):	Obj.721
Level 2 (2 elements):	Obj.76; Obj.79
Level 1 (1 elements):	Obj.678

The levels marked in blue color are removed from the base dataset to formulate reduced dataset.



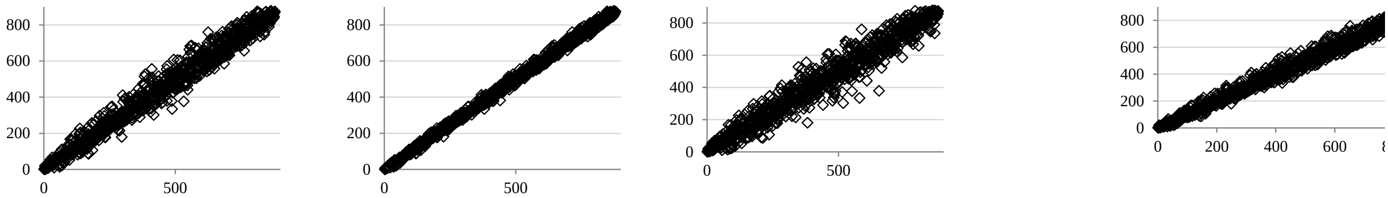
Personal Consumption Data of 1281 Danish Residents (only 1000 selected for present study)
Estimation Endpoints Using LCA (ReCiPe 2008 Method)
Endpoints Dataset (n = 1000)
Dominance Analysis Using Hasse Diagram Technique
Removal of Dominating Respondents from the Datasets
Reduced Dataset (n = 879)
Estimation of Range of Variation of Endpoints (Max and Min Values)
Generation of Dataset of Endpoints Using Random Numbers Within the Range
Random Dataset (n = 879)
Perspectives (H, I, E), Equal and Extreme Weight Scenarios
Estimation of Single Score Using LWS Method
Vector Normalization
Perspectives (H, I, E), Equal and Extreme Weight Scenarios
Estimation of Single Score Using Distance Based Method (TOPSIS)
Ranking of Respondents According to the Single Scores
Rank Correlation Analysis Using Kendall's Tau

Base dataset (n = 1000)



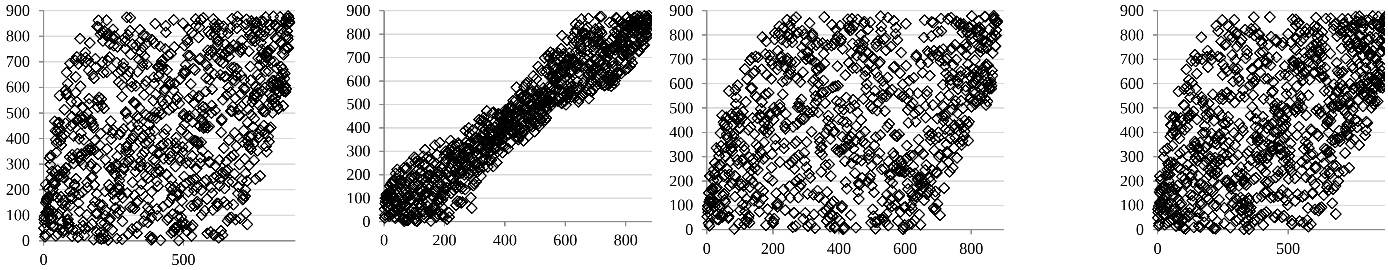
Hierarchist (0.92) Individualist (0.98) Egalitarian (0.89) Equal Weights (0.94)

Reduced Dataset after Dominance Analysis (n=879)

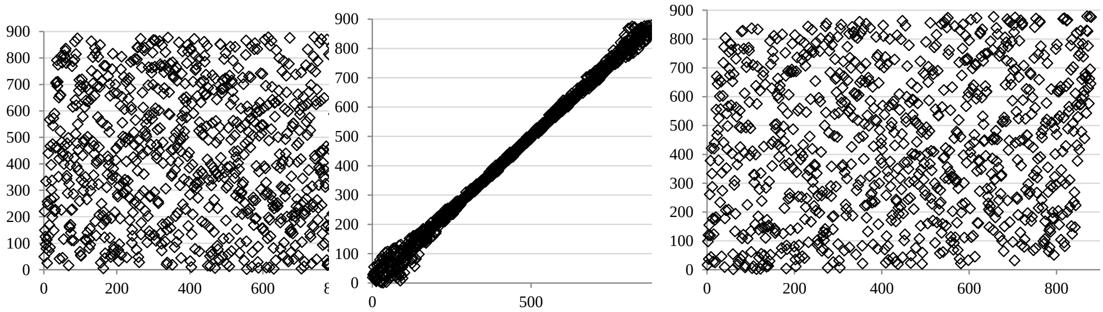


Hierarchist (0.90) Individualist (0.96) Egalitarian (0.86) Equal Weights (0.92)

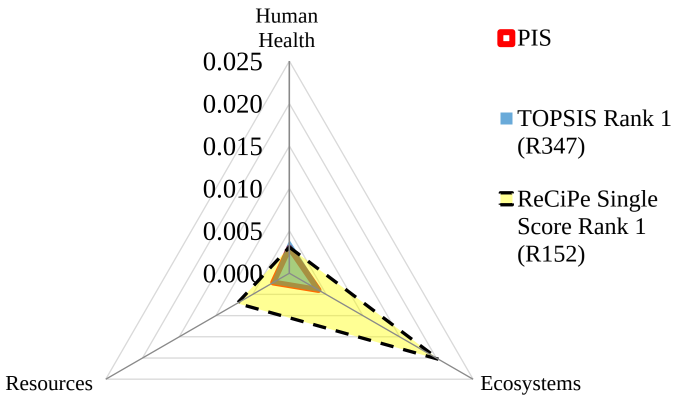
Data generated from random numbers (n = 879)



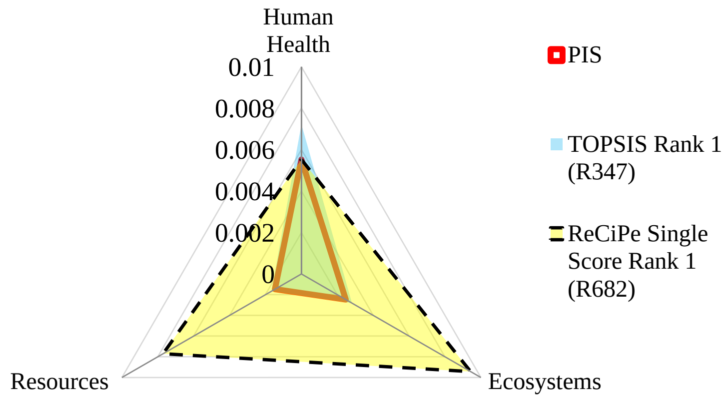
Hierarchist (0.28) Individualist (0.79) Egalitarian (0.24) Equal Weights (0.37)



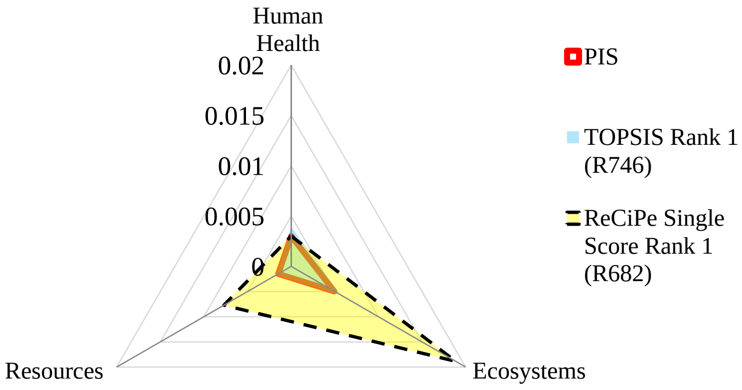
High weights to Ecosystem (0.01) High Weight to Human Health (0.96) High Weights to Resources (0.09)



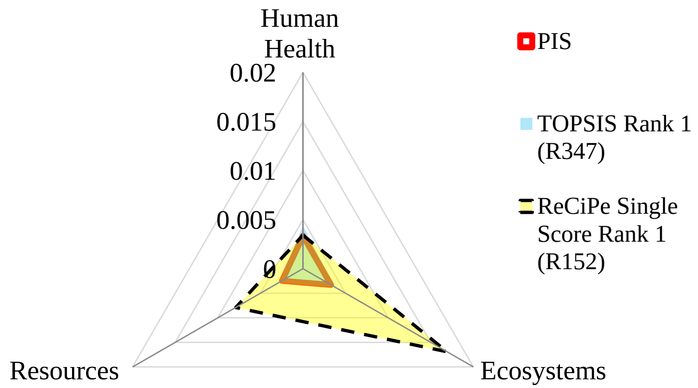
Random Dataset (*n* = 879) (Hierarchist)



Random Dataset (*n* = 879) (Individualist)



Random Dataset (*n* = 879) (Egalitarian)



Random Dataset (*n* = 879) (Equal Weights)

Weighting and Aggregation in LCA: Do Present Aggregated Single Scores Provide Correct Decision Support?

Pradip P. Kalbar, Morten Birkved, Simon Elsborg Nygaard, and Michael Hauschild

Address correspondence to:

Pradip P. Kalbar
Quantitative Sustainability Assessment Division
Dept. of Management Engineering
Technical University of Denmark (DTU)
Produktionstorvet
Building 424, room 231
2800 Kgs. Lyngby
Denmark
Tel. No.+45 45254607
Email Address: kalbar@dtu.dk; pradipkalbar@gmail.com

Summary:

This study investigates the prevailing practice of obtaining single scores in Life Cycle Assessment (LCA) and identifies potential lacunas in impact assessment methodology related to the results of aggregation into endpoints and single scores. In order to conduct this investigation, a detailed approach was adopted to facilitate identification of three main problems related to the single score calculation approach. The prevailing ReCiPe single score calculation method does not account for either the effect of so-called dominating alternatives (i.e., alternatives having high values across all endpoints) or the interdependency of the indicators being aggregated. It was also found that the simple Linear Weighted Sum (LWS) method, presently used for obtaining single scores, is not capable of accounting for the effect of weighing schemes and thus cannot realistically represent stakeholders' perspectives.

Finally, we propose a distance-based Multiple Attribute Decision Making (MADM) method for use in obtaining single scores. This method was found to be more suitable, as it takes into account the weighing schemes and types of indicators in the process of estimating single scores. The new single score calculation method proposed here is considered ideal for environmental decision-making problems in the context of Life Cycle Sustainability Assessment (LCSA). Thus, it is also ideal for situations in which more complex decision-making situations will emerge by combining LCA indicators (midpoints or endpoints) with other indicators representing the performance of a system from economic and social perspectives.

Key words: Life cycle assessment; Multiple attribute decision making; Single scores; TOPSIS; Life cycle sustainability assessment; Multiple criteria decision making

<heading level 1> Introduction

Life Cycle Assessment (LCA) of products, services and technologies has gained ever wider acceptance over the last two decades (UNEP-LCI 2012). Continuous developments in LCA have supported and strengthened the wider acceptance of LCA-based decision making in the policy arena. Newer Life Cycle Impact Assessment (LCIA) methods have been introduced; these are capable of representing the results of an LCA in the form of several non-normalized and un-weighted but still aggregated indicators (so-called endpoints). Increased use of these assessments is most likely due to the popularity of LCIA results, which facilitates easier communication. These endpoints can be further normalized (usually by external normalization) and weighted in order to obtain overall environmental performance indicators in the form of one dimensionless single indicator (a so-called single score). Just like the endpoints and for the same reasons, these single scores are (regardless of the fact that weighted results are not recommended for public dissemination by the ISO 14044:2006) becoming more popular, at least for comparative assessments (Corona et al. 2015). As with absolute assessment, it is difficult to draw any detailed conclusions from these single scores. Considering the lack of detailed information provided by single scores, as well as other possibilities of unintended uses of single scores, ISO 14044:2006 recommends providing characterized and/or normalized results along with the single score results (ISO 2006). This recommendation enables the receiving stakeholders of the LCA to judge the validity of the simplified picture provided by the single scores.

Despite the risks of over- or even misinterpreting normalized and weighted results, the demand for policy-making based on LCA, and hence simplified communication, is increasing (Hellweg and Milà i Canals 2014). Thus, normalization and weighing are becoming essential (rather than optional) parts of LCA practice (Kim et al. 2013; Van Hoof et al. 2013; Kägi et al. 2016). Many approaches to weighing the results of LCA (on midpoint as well as endpoint levels) are available. The most commonly used principles for weighing include valuation of impacts/damages in monetary terms using willingness-to-pay as a base reference, valuation of damages into costs, midpoint impacts weighing (e.g., BEES, TRACI), the Distance-to-Target approach (e.g., EDIP 97, Ecological Scarcity Method) and panel weighing (e.g., Ecoindicators 99, ReCiPe 2008). Detailed descriptions of various weighing approaches, along with their respective normalization requirements, are presented by Huppes and Oers (2011), Ahlroth et al. (2011), and Huppes et al. (2012).

The number of articles available on the weighing of LCIA results suggests that considerable effort has been spent on methods of weighing results at both the midpoint and endpoint levels. Despite the number of scientific publications on this topic, little attention has been paid to the actual aggregation procedures of the weighted impacts, with the aim of representing impacts in the form of single scores. Norris (2001) presented the problems related to single scores obtained by applying internal normalization and discussed the need for congruence in normalization methods. Seppälä, Basson, and Norris (2002) proposed a comprehensive analytical framework for LCIA and further stressed that little attention has been paid to whether or not the applied aggregation functions for obtaining single scores are appropriate. In addition, Seppälä, Basson, and Norris (2002) have noted a need to verify interdependencies among the impact category indicators, which may influence the aggregation procedure. Rogers and Seager (2009) presented a decision problem involving 5 fuel alternatives evaluated using six mid-point indicators; they applied different weighting schemes in combination with a stochastic multiple criteria evaluation method. The results obtained by Rogers and Seager (2009) revealed that there is no change in ranking results despite using different weighting schemes in accordance with the traditional LCA approach, whereas ranking results obtained by the stochastic multiple criteria evaluation method were sensitive to different weighting schemes.

The lack of attention paid to the aggregation algorithm used in obtaining single scores in conventional LCIA and the fact that single score based ranking results appear to be resistant to and independent from the weighting schemes applied provide the starting point for this study. Our study therefore investigates the need for methodological changes in the algorithms applied when computing single scores in LCIA methodology. This study thus aims at analyzing the existing calculation procedures applied in relation to single score quantification in LCIA and shows the lack of decision selectivity in present single score methods. Furthermore, our study seeks to illustrate some of the methodological lacunas in the present/prevailing practice of obtaining single scores by using a novel approach based on rank correlation analysis applied to unique respondent data. In addition, a new method for quantification of singles scores is proposed, along with detailed illustrations of how this new method performs.

<heading level 1>Methodology

Figure 1 presents this article's approach to the comparison of methodologies for obtaining single scores. As presented in Figure 1, the methodology we apply consists of four major sequences. First, life cycle impacts in the form of endpoints were estimated, thereby generating the base dataset (*n* =1000). Subsequently, single scores were estimated using the Linear Weighted Sum (LWS) method (ReCiPe method) and distance-based method. A dominance analysis was then carried out on the base data in order to identify the dominating data points. Finally, a random dataset was generated and used to identify disagreements in the aggregation methods applied to obtain single scores. Rank correlation analysis was subsequently used to investigate agreements/disagreements in the rankings. These individual steps are further explained and elaborated in the following sections.

<Figure 1 somewhere here>

<heading level 2> Estimation of life cycle impacts and single scores using the ReCiPe Method

The present study is based on the results from the Personal Metabolism (PM) – LCA model presented and described in Kalbar et al. (2016). The PM-LCA model in Kalbar et al. (2016) was applied to assess the life cycle impacts of resource consumption of urban Danish residents. All the details about how the dataset for the study was derived are presented in Supplementary Information (SI) I.

In practice, a ReCiPe Single Score is obtained using a LWS method. Assume there are *m* respondents to the above mentioned consumption questionnaire (designated as alternatives that are to be ranked, $A_1, A_2, \ldots\ldots A_m$), *n* number of attributes/environmental indicators (in our case endpoints) ($j_1, j_2, \ldots\ldots\ldots, j_n$), $w = \{w_1, w_2, \ldots\ldots, w_n\}$ are the weights assigned to each of the endpoint indicators, and x_{ij} are the normalized endpoints (external normalization) of the i^{th} respondent about the j^{th} indicator. The decision matrix, along with the weight matrix, can be represented as follows:

Attributes (endpoint indicators) $\longrightarrow$

Alternatives (Respondents) $\downarrow$

$$
\begin{array}{cccccccc}
a & X_1 & X_2 & X_3 & \cdots & X_j & \cdots & X_n \\
A_1 & x_1(a_1) & x_2(a_1) & x_3(a_1) & \cdots & x_j(a_1) & \cdots & x_n(a_1) \\
A_2 & x_1(a_2) & x_2(a_2) & x_3(a_2) & \cdots & x_j(a_2) & \cdots & x_n(a_2) \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
A_i & x_1(a_i) & x_2(a_i) & x_3(a_i) & \cdots & x_j(a_i) & \cdots & x_n(a_i) \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
A_m & x_1(a_m) & x_1(a_m) & x_1(a_m) & \cdots & x_j(a_m) & \cdots & x_n(a_m) \\
\cancel{a} \rightarrow & w_1 & w_2 & w_3 & \cdots & w_j & \cdots & w_n
\end{array}
$$

Weights

In the present study, there are 1000 respondents and 3 endpoint indicators; hence, the decision matrix size was 1000 x 3. After the formulation of the decision matrix, the respondent with the lowest single score (A^*) can be identified from following equation.

$$A^* = \left\{ A_i \middle| \min_{i} \sum_{j=1}^{n} w_j x_{ij} \right\} \qquad \text{for } i=1,2,3,\ldots\ldots,m \qquad (1)$$

Relying on the results obtained from Eq. [1], all respondents were ranked in ascending order, according to the magnitude of their single scores, in such a way that those respondents with lower single scores ranked first (A*) and respondents with higher single score ranked last. The ranks obtained were subsequently used for the rank correlation analysis.

<heading level 2> Distance-based approach for estimation of single scores

Multiple Attribute Decision-Making (MADM) consists of a multitude of methods for the evaluation of alternatives based on indicators/attributes. Utility based methods (i.e., linear sum method, LWS), outranking methods providing partial/complete rankings (ELECTREE, PROMETHEE, Hasse Diagram Technique), distance-based methods (compromise programming, TOPSIS) are some of the most commonly used methods in MADM (Hwang and Yoon 1981; Yoon and Hwang 1995; Pohekar and Ramachandran 2004; Kiker et al. 2005; Figueira, Greco, and Ehrgott 2005; Behzadian et al. 2012).

In the present study, a distance-based MADM method, the Technique for Order Preference by Similarity to Ideal Solutions (TOPSIS), was used to estimate single scores using endpoints obtained from the ReCiPe method. TOPSIS was selected because it is one of the most widely used methods, is easy to use and implement, mimics human thinking (Behzadian et al. 2012) and has proven to have the lowest rank reversal (change in ranks by addition/deletion of alternative) compared to similar methods (Zanakis et al. 1998; Shih, Shyur, and Lee 2007; Kalbar, Karmakar, and Asolekar 2012; Kalbar, Karmakar, and Asolekar 2015; Kim, Park, and Yoon 1997). The TOPSIS method chooses the alternative that is nearest to the formulated ideal solution and farthest from the formulated non-ideal solution. The ideal and non-ideal solutions are defined based on the type of attribute (cost or benefit type) and can thus handle multidimensional problems (Kalbar, Karmakar, and Asolekar 2012).

Following the notations used earlier for alternatives, indicators and weights, x_{ij} is the normalized endpoint (i.e., the vector representing externally normalized endpoint) of the i^{th} respondent about the j^{th} indicator. The matrix (x_{ij}) is further vector normalized using the following equation.

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{m} x_{ij}^2}} \tag{2}$$

The internally normalized endpoint matrix (r_{ij}) is then multiplied by the weight matrix (w_j) to obtain the weighted normalized endpoint matrix (v_{ij})

$$v_{ij} = w_j r_{ij} \tag{3}$$

The positive ideal solution, labeled *PIS*, (v_{ij}^{+}) and the non-ideal solution labeled *NIS*, (v_{ij}^{-}) can then be formulated using the following equations.

$$PIS, \quad v_{ij}^{+} = \{v_1^{+}, v_2^{+}, \dots \dots .., v_j^{+}, \dots \dots ., v_n^{+}\}$$

$$= \left\{\left(\begin{matrix} max \\ i \end{matrix} v_{ij} \middle| j \in J_1\right), \left(\begin{matrix} min \\ i \end{matrix} v_{ij} \middle| j \in J_2\right) \middle| i = 1, \dots \dots \dots ., m\right\} \tag{4}$$

$$NIS, \quad v_{ij}^{-} = \{v_1^{-}, v_2^{-}, \dots \dots .., v_j^{-}, \dots \dots ., v_n^{-}\}$$

$$= \left\{\left(\begin{matrix} min \\ i \end{matrix} v_{ij} \middle| j \in J_1\right), \left(\begin{matrix} max \\ i \end{matrix} v_{ij} \middle| j \in J_2\right) \middle| i = 1, \dots \dots \dots ., m\right\} \tag{5}$$

where J_1 is a set of benefit type attributes (or indicators), J_2 is a set of cost type attributes, and $J_1 + J_2 = n$, *i.e.,* the total number of attributes. Benefit type indicators are indicators that represent monotonic utilities, i.e., the greater the indicator value, the more it is preferred (e.g., fuel efficiency, production yield). In contrast, cost type indicators are indicators representing decreasing monotonic utility, i.e., the greater the indicator value, the less it is preferred (e.g., production cost, environmental impact indicators).

In the present study, there are three endpoint indicators, all three of them are cost type indicators (as they represent damages, i.e., loss of value, damage to the environment), and all three hence belong to set J_2.

Now the distance of each alternative to the formulated ideal and non-ideal solutions can be estimated as:

Distance to ideal solution, $$D_i^+ = \sqrt{\sum_{j=1}^{n}\left(v_{ij} - v_j^+\right)^2}\ , i=1,2,\ldots\ldots\ldots,m \tag{6}$$

Distance to non-ideal solution, $$D_i^- = \sqrt{\sum_{j=1}^{n}\left(v_{ij} - v_j^-\right)^2}\ , i=1,2,\ldots\ldots\ldots,m \tag{7}$$

Finally, the single score calculated for each respondent, in terms of relative closeness (C_i^*) to the ideal and non-ideal solutions can be estimated as:

$$C_i^* = \frac{D_i^-}{\left(D_i^+ + D_i^-\right)}, i=1,2,\ldots\ldots\ldots,m \tag{8}$$

From equation (8) it can be seen that the value of C_i^* ranges between 0 and 1.

Using the procedure presented in equations (4-8), closeness to the ideal/non-ideal solutions was estimated for all respondents. Subsequently, the respondents were ranked in descending order according to C_i* in such a way that the highest relative closeness to the ideal solution (*i.e.*, nearest to ideal solution and farthest from non-ideal solution) had the highest rank while the lowest relative closeness (*i.e.*, farthest from ideal solution and nearest to non-ideal solution) had the lowest rank. Using this procedure, the ranks for the different cultural perspectives and scenarios (pls. see Table 1) were calculated. These ranks were later used for rank correlation analyses.

<heading level 2> Evaluation of correlation among datasets and ranks

To analyze the correlation among the datasets (endpoints), Spearman's rank correlation coefficient (ρ) was used. Non-parametric rank correlation analysis was chosen due to the robustness of the rank correlation towards data outliers (Reimann et al. 2011). As mentioned earlier, the respondents were ranked according to the scores obtained by applying the above-mentioned methods of ReCiPe and TOPSIS. For the assessment of the correlation structures among the ranks obtained from the LWS method applied in ReCiPe and the distance-based method applied in TOPSIS for all of the seven scenarios presented in Table S1 in SI I, Kendall's rank correlation coefficient (τ) was used.

<heading level 1> Results

Upon assessment of the effects of the aggregation method used for obtaining single scores, both the original and reduced datasets were found unsuitable as they were strongly correlated and further dominating respondents were identified in the original dataset. Hence, all the results and discussions relating to the original and reduced dataset are presented in SI I.

A random data set is needed to evaluate the decision making process based on single scores by rank analysis. This is because a dataset with correlation among the indicators or dominating respondents is not suitable to demonstrate the difference in decision making between the different approaches (LWS and TOPSIS), as illustrated in the SI I. Maximum and minimum values for the three endpoints derived from the reduced data (refer to SI I) were identified, thereby enabling the formulation of the endpoint ranges. Random numbers were subsequently generated, respecting the identified ranges. Figure S5 in SI I shows the boxplots of the random numbers generated. The correlation between the endpoints generated using random numbers was also analyzed (see Table S3 in SI I). This correlation analysis revealed that the random endpoints are not correlated (ρ = 0.02).

The random endpoint dataset was assumed to be a good surrogate endpoint dataset (i.e., a random dataset with independent endpoints) and usable for obtaining single scores (based on independent endpoints using two single score calculation methods).

The weighted normalized endpoint values of the best performing respondents are compared with PIS and shown in Figure 2. The PIS illustrated with red line triangles is the best possible environmental profile from the dataset; the objective of the methods used for obtaining single scores is to match the shape of the triangle formed by the PIS. It is evident from the correlation analysis that there is a complete disagreement between the results obtained by ReCiPe and TOPSIS (pls. refer to Figure 3c and Table 1).

At this point, it is important to note that the distance-based method TOPSIS is capable of consistently identifying the respondent closest to PIS (as it can be seen from Figure 2 TOPSIS identifies the best performing respondents almost identically with the triangle formed by PIS in each of the perspectives, as well as in the equal weight scenario). This improved selectivity is further underlined by the rank correlation results obtained for the entire random dataset and shown in Table 1. To further investigate and validate the performance of the two aggregation methods, the extreme weight scenarios were used. The results of the extreme weight scenarios are shown in Figure S6 and Table S5 in SI I. These results show that there is a disagreement between the ranks obtained by the TOPSIS method for all three cultural perspectives as well as for different weighting schemes given in Table S1 in SI I. There is in addition strong agreement between the ranks generated by the ReCiPe single score method for all the three cultural perspectives as well as for the equal and extreme weights scenarios.

<Figure 2 somewhere here>

<Figure 3 somewhere here>

<Table 1 somewhere here>

<heading level 1> Discussion

Our study systematically investigated methodological issues related to the use of single scores obtained from a contemporary impact assessment method often used in LCA. The ReCiPe single score was specifically chosen for our evaluation because it is contemporary, widely used and well recognized. Three sets of data were used to demonstrate the need for improvement to the present practice of aggregating endpoints and interpreting single scores. Three major issues relating to the decision support provided by the ReCiPe single scores were identified and are discussed in the following subsections.

<heading level 2> Presence of dominating alternatives

The first major finding was that the sample of alternatives under evaluation may contain one or more dominating alternatives (i.e., in our case, respondents with considerably higher/lower values than neighboring and average respondents across all three endpoints). This finding became evident from the results of the base dataset (n = 1000), where both methods applied for the purpose of obtaining single scores identified the same respondent (R678) as best performing, regardless of the cultural perspective applied or the weighting scenario used (including equal weights). Thus, we conclude that the presence of dominating alternatives masks the decision dependency of the weighting scheme applied in accordance with each of the perspectives and scenarios. None of the methods would help to identify different respondents for different perspectives.

Rogers and Seager (2009) reported the same problem of insensitivity towards weighting schemes. However, they applied yet another MADM method for comparison of fuel alternatives at the midpoint level. Rogers and Seager (2009) concluded that the insensitivity problem occurred due to bias introduced via the external normalization practice of LCA. The fact that midpoint normalization may introduce considerable high bias is well documented in the literature and thus internal normalization is recommended to minimize the effect of bias (White and Carty 2010; Curran 2012).

Our study focuses on respondent comparison at the endpoint level. Endpoint normalization is more stable and introduces less bias, as only a few midpoints contribute to a given endpoint and thus bias due to incompleteness of data has a lower impact at the endpoint level (Van Hoof et al. 2013). In addition, as seen in Figure 1, more specifically in the TOPSIS application approach, the externally normalized endpoints were, in addition, internally normalized (i.e., using vector normalization). Hence, the sole reason for identifying the same alternative as the best performing respondent, regardless of the cultural perspective, was the presence of dominating alternatives.

<heading level 2> Correlation/dependence among endpoints

After removing the dominating alternatives, a new (reduced) set of endpoint results was obtained for the respondents (see Figure S3 in SI I). Within the results obtained for this reduced dataset, there was still no observable effect on the ranks that were obtained by the two ranking methods applying different weighting schemes. It was also found that a strong correlation between the endpoints (see Table S3 in SI I) was the only factor responsible for the observed insensitivity towards different weighting schemes. This insensitivity suggests that the endpoints are preferentially dependent. As discussed by (Seppälä, Basson, and Norris 2002), if the attributes are preferentially dependent then the application of linear weighted methods for obtaining single scores (which are based on the utility approach) will fail. The preferential dependencies restrain simple aggregation methods (such as linear methods) from taking into account the effects of different weighting schemes. Hence, there is a need to investigate other approaches to the aggregation of impact assessment results in LCA. Our study pursues precisely this target by applying a distance-based approach to obtaining single scores. As seen from the results (in Figure S4 in SI I), the TOPSIS approach does perform better than the ReCiPe single score approach by at least identifying different respondents as best performing for various cultural perspectives and different weighting schemes.

<heading level 2> Need for change in present aggregation methods

To further investigate the lack of decision selectivity in terms of weighting schemes' sensitivity to simple result aggregation approaches, we had to use a surrogate dataset consisting of endpoints generated using random numbers. This is because it is not possible to obtain endpoints based on characterization that are totally independent or just exhibit low correlations. The results of the culturally specific ranks obtained for the random dataset (see Figure 2 and Table 1) reveals that the TOPSIS method performs better than the LWS, because of the distance-based mathematical approach incorporated in TOPSIS.

Similarly, the results of the extreme weight scenarios indicate that (see Figure S6, Table S5 in SI II and Figure 3c) TOPSIS can generate ranks that reflect the weights of the individual cultural perspectives and scenarios (even extreme weights). It is evident from the results obtained for the random dataset that the ReCiPe single score-based ranks are insensitive to different weighting schemes. This insensitivity is not only limited to ReCiPe single score but to all the LCIA methods using simple aggregation method such as LWS. The reason for the insensitivity towards different endpoint weighting schemes lies in the mathematical approach (which is utility based), followed by the linear weighting method used for obtaining single scores. This mathematical approach yields a preference for alternatives (in our case respondents) that are independent of the stakeholder's value choices (which are embedded in the weighting schemes). Similar results relating to the disadvantages of LWS aggregation approaches were presented by Norris (2001). Amine, Pailhes, and Perry (2014) evaluated five MADM methods (including a LWS method and a weighted product method) and concluded that the TOPSIS approach generated more consistent ranks, capable of reflecting the decision maker's preferences and thus value choices (i.e., weighting schemes). The reported superior decision selectivity of TOPSIS aligns with the findings of our study, where the impacts of different weighting schemes (and hence value choices) are clearly seen in the TOPSIS results and not in the LWS method derived results.

The problems related to the LWS approach become severe when conflicting attributes are involved, such as benefit or cost type attributes. The basic principles of additive utility are violated when dealing with decision problems with multiple dimensions by applying an LWS approach (Pohekar and Ramachandran 2004). Moreover, other simple utility functions (multiplicative) will also entail the exact same disadvantages as LWS. The TOPSIS method includes an inherent mechanism to handle both best and cost attributes effectively (see Eqs. 3 and 4) and hence performs better than most of the other MADM methods.

Apart from the above-mentioned advantages of the TOPSIS method, another advantage it offers is that its results can be more clearly presented in terms of weighted normalized values of the indicators. These can then be compared with PIS and/or NIS solutions, which are the theoretical best and worst possible alternatives (i.e., benchmarks). This comparative presentation also provides an idea of the best/worst possible achievable targets identified by the TOPSIS method, which can be used for benchmarking the environmental performance of products and services.

<heading level 1> Conclusions

The problems with the conventional approaches for obtaining single scores in LCA show that there is a need to change the methods used for aggregating endpoints (and mid-points). More realistic endpoints (with lower inter-dependencies) will most likely emerge in the near future, with the development of more complete models for life cycle inventory analysis and with LCIA covering far more emissions and end-of-life scenarios. In this study, we have systematically demonstrated the effectiveness of one such proposed new method, TOPSIS, for obtaining single scores.

In the context of LCSA and as reported in Guinée et al. (2011), more complex decision situations are expected to emerge (i.e., the more (conflicting) indicators, the more complex the decision situation). These decision situations will normally involve LCA indicators (midpoints or endpoints) along with other indicators representing the performance of the system from economic and social perspectives. These indicators will have different units (as they are derived from different tools and techniques) and different types (cost types such as endpoints/capital costs or benefit types such as social indicators representing acceptability of the products/services). With such an indicator set representing complex decision-making situations, it is essential to use more efficient multi-criteria decision-making methods such as TOPSIS.

The adoption of new methods for obtaining single scores will provide more rational decision support in terms of accounting for the positive and negative aspects of products/services. This unique approach to obtaining single scores will highlight the alternative that best matches with the theoretically positive ideal solution, which will be at the same time worst matching with the theoretically negative ideal solution. In LCSA, where there are definitely conflicting indicators, methods such as TOPSIS will play a critical role in prioritizing alternatives in the context of overall sustainability.

Acknowledgements:

The first author acknowledges Postdoctoral fellowship received from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement no 609405 (COFUNDPostdocDTU). Authors also acknowledge insightful comments received from three anonymous reviewers. Authors do not have any conflict of interest.

About the Authors:

Pradip Kalbar is Postdoc, Morten Birkved is Associate Professor, and Michael Hauschild is Professor at Quantitative Sustainability Assessment Division at the Dept. of Management Engineering, Technical University of Denmark in Kongens Lyngby, Denmark. Simon Elsborg Nygaard is Ph.D. Candidate at the Department of Psychology and Behavioral Sciences, Aarhus University, Denmark

References:

Ahlroth, Sofia, Måns Nilsson, Göran Finnveden, Olof Hjelm, and Elisabeth Hochschorner. 2011. “Weighting and Valuation in Selected Environmental Systems Analysis Tools - Suggestions for Further Developments.” Journal of Cleaner Production 19 (2-3): 145–56. doi:10.1016/j.jclepro.2010.04.016.

Amine, Mehdi El, Jérôme Pailhes, and Nicolas Perry. 2014. “Critical Review of Multi-Criteria Decision Aid Methods in Conceptual Design Phases: Application to the Development of a Solar Collector Structure.” Procedia CIRP 21. Elsevier B.V.: 497–502. doi:10.1016/j.procir.2014.03.134.

Behzadian, Majid, S. Khanmohammadi Otaghsara, Morteza Yazdani, and Joshua Ignatius. 2012. “A State-of the-Art Survey of TOPSIS Applications.” Expert Systems with Applications 39 (17). Elsevier Ltd: 13051–69. doi:10.1016/j.eswa.2012.05.056.

Corona, A, C M Markussen, M Birkved, and B Madsen. 2015. “Comparative Environmental Sustainability Assessment of Bio-Based Fibre Reinforcement Materials for Wind Turbine Blades.” Wind Engineering 39 (1): 53–64. doi:10.1260/0309-524X.39.1.53.

Curran, Mary Ann. 2012. Life Cycle Assessment Handbook: A Guide for Environmentally Sustainable Products. John Wiley & Sons.

Figueira, José, Salvatore Greco, and Matthias Ehrgott. 2005. Multiple Criteria Decision Analysis: State of the Art Surveys. Vol. 78. Springer Science & Business Media.

Guinée, Jeroen B, Reinout Heijungs, Gjalt Huppes, Alessandra Zamagni, Paolo Masoni, Roberto Buonamici, Tomas Ekvall, and Tomas Rydberg. 2011. “Life Cycle Assessment: Past, Present, and Future.” Environmental Science & Technology 45 (1): 90–96. doi:10.1021/es101316v.

Hellweg, Stefanie, and Llorenç Milà i Canals. 2014. “Emerging Approaches, Challenges and Opportunities in Life Cycle Assessment.” Science (New York, N.Y.) 344 (6188): 1109–13. doi:10.1126/science.1248361.

Huppes, Gjalt, and Lauran Van Oers. 2011. Background Review of Existing Weighting Approaches in Life Cycle Impact Assessment (LCIA). doi:10.2788/88828.

Huppes, Gjalt, Lauran Van Oers, Ugo Pretato, and David W. Pennington. 2012. “Weighting Environmental Effects: Analytic Survey with Operational Evaluation Methods and a

Meta-Method." International Journal of Life Cycle Assessment 17 (7): 876–91. doi:10.1007/s11367-012-0415-x.

Hwang, Ching-Lai, and K.Paul Yoon. 1981. Multiple Attribute Decision Making. Springer-Verlag, Berlin.

ISO. 2006. ISO 14044:2006 Environmental Management—life Cycle Assessment—requirements and Guidelines.

Kägi, Thomas, Fredy Dinkel, Rolf Frischknecht, Sebastien Humbert, Jacob Lindberg, Steven De Mester, Tommie Ponsioen, Serenella Sala, and Urs Walter Schenker. 2016. "Session 'Midpoint, Endpoint or Single Score for Decision-making?'—SETAC Europe 25th Annual Meeting, May 5th, 2015." International Journal of Life Cycle Assessment 21 (1): 129–32. doi:10.1007/s11367-015-0998-0.

Kalbar, Pradip P., Morten Birkved, Simon Kabins, and Simon E Nygaard. 2016. "Personal-Metabolism (PM) Coupled with Life Cycle Assessment (LCA) Model: Danish Case Study." Environment International 91:168-179. doi:10.1016/j.envint.2016.02.032.

Kalbar, Pradip P., Subhankar Karmakar, and Asolekar S. R. 2012. "Selection of an Appropriate Wastewater Treatment Technology: A Scenario-Based Multiple-Attribute Decision-Making Approach." Journal of Environmental Management 113. 158–69. doi:10.1016/j.jenvman.2012.08.025.

Kalbar, Pradip P., Subhankar Karmakar, and Asolekar S. R. 2015. "Selection of Wastewater Treatment Alternative: Significance of Choosing MADM." Environmental Engineering & Management Journal 14 (5): 1011–1120.

Kiker, Gregory a, Todd S Bridges, Arun Varghese, P Thomas P Seager, and Igor Linkov. 2005. "Application of Multicriteria Decision Analysis in Environmental Decision Making." Integrated Environmental Assessment and Management 1 (2): 95–108. doi:10.1897/IEAM_2004a-015.1.

Kim, Gyutai, Chan S Park, and K.Paul Yoon. 1997. "Identifying Investment Opportunities for Advanced Manufacturing Systems with Comparative-Integrated Performance Measurement." International Journal of Production Economics 50 (1): 23–33. doi:10.1016/S0925-5273(97)00014-5.

Kim, Junbeum, Yi Yang, Junghan Bae, and Sangwon Suh. 2013. “The Importance of Normalization References in Interpreting Life Cycle Assessment Results.” Journal of Industrial Ecology 17 (3): 385–95. doi:10.1111/j.1530-9290.2012.00535.x.

Norris, Gregory. 2001. “The Requirement for Congruence in Normalization.” The International Journal of Life Cycle Assessment 6 (2): 85–88. doi:10.1007/BF02977843.

Pohekar, S. D., and M. Ramachandran. 2004. “Application of Multi-Criteria Decision Making to Sustainable Energy Planning - A Review.” Renewable and Sustainable Energy Reviews 8: 365–81. doi:10.1016/j.rser.2003.12.007.

Reimann, Clemens, Peter Filzmoser, Robert Garrett, and Rudolf Dutter. 2011. Statistical Data Analysis Explained: Applied Environmental Statistics with R. John Wiley & Sons.

Rogers, Kristin, and Thomas P Seager. 2009. “Environmental Decision-Making Using Life Cycle Impact Assessment and Stochastic Multiattribute Decision Analysis: A Case Study on Alternative Transportation Fuels.” Environmental Science & Technology 43 (6): 1718–23. doi:10.1021/es801123h.

Seppälä, Jyri, Lauren Basson, and Gregory a Norris. 2002. “Decision Analysis Frameworks for Life-Cycle Impact Assessment.” Journal of Industrial Ecology 5 (4): 45–68. doi:10.1162/10881980160084033.

Shih, Hsu Shih, Huan Jyh Shyur, and E. Stanley Lee. 2007. “An Extension of TOPSIS for Group Decision Making.” Mathematical and Computer Modelling 45: 801–13. doi:10.1016/j.mcm.2006.03.023.

UNEP-LCI. 2012. Greening the Economy through Life Cycle Thinking: Ten Years of the UNEP/SETAC Life Cycle Initiative. United Nations Environment Programmme, 2012.

Van Hoof, Gert, Marisa Vieira, Maria Gausman, and Annie Weisbrod. 2013. “Indicator Selection in Life Cycle Assessment to Enable Decision Making: Issues and Solutions.” International Journal of Life Cycle Assessment 18 (8): 1568–80. doi:10.1007/s11367-013-0595-z.

White, Philip, and Mark Carty. 2010. “Reducing Bias through Process Inventory Dataset Normalization.” International Journal of Life Cycle Assessment 15 (9): 994–1013. doi:10.1007/s11367-010-0215-0.

Yoon, K Paul, and Ching-Lai Hwang. 1995. Multiple Attribute Decision Making: An Introduction. Vol. 104. Sage publications.

Zanakis, Stelios H., Anthony Solomon, Nicole Wishart, and Sandipa Dublish. 1998. "Multi-Attribute Decision Making: A Simulation Comparison of Select Methods." European Journal of Operational Research 107 (3): 507–29. doi:10.1016/S0377-2217(97)00147-1.

Figure Captions:

Figure 1: The assessment sequences followed for comparison of single score quantification approaches

Figure 2: The radar plot shows the weighted normalized values of endpoints for best performing respondents of the random dataset for various cultural perspectives and the equal weight scenario (a) shows Hierarchist, (b) shows Individualist, (c) shows Ealitarian and (d) shows Equal Weights

Figure 3: Graphs showing the correlation of ranks generated by the TOPSIS method (x-axis) and the ranks generated by the Linear Weighted Method - ReCiPe Single Score (y-axis). (a) presents the correlation analysis of the base dataset (n =1000), (b) presents the correlation analysis for the reduced dataset (n = 879), while (c) presents the correlation results for the random dataset (n =879). The values in parentheses are the Kendall's rank correlation coefficient (t) values obtained for each of the analyses.

Table Captions:

Table 1: Results of Kendall's rank correlation coefficient (τ) between the ranks generated by the two methods for various perspectives (Random Dataset)